

大数据职称考试知识大纲

大数据分析应用-初级

2024 年版

目录

第一部分 基础知识	1
一、大数据法律法规、政策文件、相关标准	1
二、计算机基础知识	2
三、信息化基础知识	3
四、密码学	5
五、大数据安全	6
六、数据库系统	6
七、数据仓库	7
第二部分 专业知识	7
一、大数据技术与应用	7
二、大数据分析模型	8
三、数据科学	8

第一部分 基础知识

一、大数据法律法规、政策文件、相关标准

（一）法律法规

- 1、《中华人民共和国民法典》中华人民共和国主席令第四十五号
(1) 了解隐私权和个人信息保护有关内容。
- 2、《中华人民共和国数据安全法》中华人民共和国主席令第八十四号
(1) 了解数据安全制度、数据安全保护义务、政务数据安全与开放等相关内容。
- 3、《中华人民共和国网络安全法》中华人民共和国主席令第五十三号
(1) 了解数据分类、重要数据备份和加密等措施。
- 4、《中华人民共和国电子签名法（2019 版）》中华人民共和国主席令第二十九号
(1) 了解电子签名、数据电文基本概念。
- 5、《中华人民共和国密码法》中华人民共和国主席令第三十五号
(1) 了解核心密码、普通密码、商用密码之间的区别。
- 6、《中华人民共和国个人信息保护法》中华人民共和国主席令第九十一号
(1) 了解敏感个人信息处理规则、非法获取个人信息相关法律责任。
- 7、《关键信息基础设施安全保护条例》中华人民共和国国务院令 745 号
(1) 了解关键信息基础设施的基本概念、认定和监管部门责任分工。
- 8、《网络安全等级保护条例》
(1) 了解等级保护适用范围、基本要求、等级划分和评估规范。
- 9、《数据出境安全评估办法》国家互联网信息办公室令 11 号
(1) 了解数据出境安全评估原则、数据安全保护责任义务。

（二）相关标准

- 1、GB/T 35295-2017 信息技术 大数据 术语
(1) 了解本标准的范围、术语和定义，以及密切相关的通用术语。
- 2、GB/T 38676-2020 信息技术 大数据 存储与处理系统功能测试要求
(1) 了解大数据、存储与处理系统术语和定义。
- 3、GB/T 42130-2022 智能制造 工业大数据系统功能要求
(1) 了解智能制造、工业大数据系统术语和定义。
(2) 了解工业大数据系统功能框架。

（三）职业道德

- 1、掌握专业技术人员职业道德的基本要求。

二、计算机基础知识

（一）面向对象技术

- 1、了解面向对象技术的基本概念、特点及方法等。
- 2、掌握面向对象编程与非面向对象编程的优缺点。
- 3、掌握抽象的 5 个层次、抽象形式。
- 4、了解类和方法。
- 5、了解消息、实例和初始化。

（二）数据结构与算法

- 1、掌握程序性能分析的概念和方法，包括时间复杂性与空间复杂性分析。
- 2、掌握线性表的概念，掌握堆栈、队列、跳表和散列的描述方法与应用。
- 3、掌握常用的排序方法与应用。
- 4、掌握树的描述方法与应用。
- 5、掌握图的描述方法与应用。

（三）操作系统

- 1、掌握操作系统的概念和操作系统结构。
- 2、掌握操作系统的进程管理，包括进程概念、进程调度、同步及死锁处理。
- 3、了解内存管理，包括内存管理策略和虚拟内存管理。
- 4、了解存储管理，包括文件系统、文件系统实现、大容量存储结构和 I/O 系统。
- 5、了解系统保护与安全。

（四）计算机网络

- 1、了解计算机网络在信息时代的核心作用以及计算机网络的发展历史，理解计算机网络的分类、网络标准化、网络参考模型、网络体系结构。
- 2、掌握物理层的基本概念，理解信道极限容量的概念以及信道最大传输速率的公式，理解模拟传输和数字化传输的物理层标准。
- 3、掌握数据链路层的基本概念，理解停等协议和连续 ARQ 协议、滑动窗口协议、检错和纠错机制。
- 4、掌握介质访问控制子层的基本概念，理解动态多路访问控制协议、以太网、无线局域网、数据链路层的交换技术。
- 5、了解网络层的基本概念，了解路由协议、拥塞控制算法、服务质量、网络互连、IP 协议、子网掩码。

- 6、了解传输层的基本概念，了解传输层路由协议、TCP、UDP、拥塞控制算法。
- 7、了解应用层的基本概念，了解 DNS、邮件系统、WWW、流音频与视频。

三、信息化基础知识

（一）数字政府

1、数字政府的概念

- （1）掌握数字政府的内涵与外延。
- （2）掌握数字政府的构成要素。
- （3）熟悉数字政府的主要技术基础。

2、数字政府数据治理

- （1）掌握数字政府数据治理的概念。
- （2）了解数据治理的现状。
- （3）掌握数据治理的主要内容和流程。

3、数字政府数据开放与共享

- （1）熟悉数字政府数据开放的概念。
- （2）了解数字政府数据共享的相关概念。

（二）电子商务

1、电子商务的概念

- （1）掌握电子商务的发展与定义。
- （2）掌握电子商务的交易过程。
- （3）熟悉电子商务的分类和主体框架。
- （4）熟悉电子商务的商业模式：B2B、B2C、C2C、O2O。

2、网络营销与管理

- （1）熟悉电子商店的构成、功能与网上购物流程。
- （2）掌握电子商店的日常运营管理。
- （3）了解网络营销与管理的相关概念：网络广告、搜索引擎营销、电子邮件营销、微博营销、微信营销、直播营销、大数据营销等。

3、电子商务规划与实施

- （1）熟悉电子商务系统设计架构与要素。
- （2）熟悉电子商务系统的建设过程。
- （3）了解电子商务支付系统的基本构成、功能与种类。

（三）新一代信息技术

1、零信任、可信网络架构

- （1）了解零信任网络、可信网络的概念。
- （2）理解零信任网络架构的理念，主要包括以身份为中心、环境感知、动态的权限控制等。
- （3）理解可信网络架构的特征。

2、移动互联网

- （1）掌握 5G 移动互联网的概念。
- （2）理解 5G 移动互联网的特点，包括高速率、低时延和大连接等。
- （3）了解 6G 的概念。

3、云计算

- （1）了解云计算的概念、架构。
- （2）熟悉云计算服务的类型。
- （3）熟悉云、边、端一体化架构及相关技术。

4、物联网

- （1）了解物联网的概念、网络架构。
- （2）理解物联网的应用及相关的安全风险。

5、人工智能

- （1）理解人工智能大模型的概念。
- （2）了解国外 ChatGPT、Sora 等人工智能模型。
- （3）熟悉国内相关的人工智能大模型。
- （3）熟悉人工智能实际应用。

6、区块链

- （1）了解区块链的概念。
- （2）理解区块链的系统架构，包括数据层、网络层、共识层、激励层、合约层和应用层。
- （3）理解区块链的公有链、联盟链及私有链的概念及安全问题。

（四）信息系统项目管理

1、信息系统项目管理基本概念

- （1）掌握信息系统治理的任务、方法与标准。
- （2）掌握信息系统审计的方法、流程与内容。
- （3）熟悉信息系统管理的基本概念。

(4) 了解项目信息系统工程的基本概念。

(5) 掌握项目管理知识体系的构成。

2、项目管理过程

(1) 掌握项目立项管理。

(2) 了解项目范围管理。

(3) 掌握项目进度管理。

(4) 掌握项目成本管理。

(5) 掌握项目质量管理。

(6) 掌握项目资源管理。

(7) 掌握项目沟通管理。

(8) 了解项目风险管理。

3、高级项目管理

(1) 掌握配置与变更管理。

(2) 掌握项目集管理。

(3) 掌握项目组合管理。

(4) 熟悉组织级项目管理。

(5) 了解量化项目管理。

四、密码学

(一) 密码学概述

1、了解密码学发展史、密码学概念。

(二) 密码学分类

1、了解密码学的分类，包括按时间、按密钥、按密码体制和按明文处理方法分类等。

(三) 加密技术

1、了解古典密码学、对称密码算法、非对称密码算法、哈希算法、散列函数(MD5算法)、数字签名技术。

2、掌握软件与硬件加密技术。

(四) PKI 概述

1、了解 PKI 的作用、体系结构。

2、掌握 PKI 的组成部分。

(1) 公钥技术 (2) 数字证书 (3) CA (颁发机构) (4) RA (注册机构)

3、掌握常见的 PKI 证书 X.509 和 PKCS(公钥加密标准)等系列标准。

(五) 数字证书

1、掌握数字证书的概念、签发流程。

2、理解 CA 系统的组成、CA 的现状及应用。

(六) 国密算法、商密算法

1、了解国密算法、商密算法的概念。

2、熟悉国密算法 SM1、SM2、SM3、SM4、SM7、SM9、祖冲之密码算法等的作用。

3、熟悉常用商密算法的作用。

五、大数据安全

(一) 网络安全

1、熟悉网络安全的一些基本概念、基本属性及网络安全模型等。

2、了解网络安全体系的安全服务、安全机制、安全服务配置和安全管理等。

3、熟悉各种网络安全技术的概念，了解物理安全、网络安全、系统安全、应用安全、网络设备安全的基本知识。

4、了解网络安全运维与保障的框架、基本原理等知识。

(二) 数据安全

1、了解大数据安全管理及安全需求的原则，掌握数据安全分类、定级的原则。

2、了解数据采集、存储、处理、分发、删除等活动的概念，了解相关安全技术与安全要求。

3、熟悉云存储安全体系、数据生命周期中的安全风险、保障云存储安全的原则及云计算边界安全相关的概念。

4、了解大数据安全与防护保障机制。

5、了解信息系统安全运维与安全管理的模型、体系、方法等基本知识，掌握信息系统安全风险评估和等级保护、信息安全应急响应的相关要求和方法。

六、数据库系统

1、掌握关系模型概念和 SQL 语言。

2、掌握关系数据库设计方法。

- 3、了解关系数据库事务概念和事务调度方法。
- 4、了解数据库并发控制技术。
- 5、了解非关系数据库背景、特点和分类。

七、数据仓库

- 1、了解数据挖掘的定义、功能、常用方法。
- 2、了解数据仓库的产生与发展，掌握数据仓库的定义。
- 3、了解数据仓库与数据挖掘的联系与区别。
- 4、了解常用数据挖掘工具。
- 5、理解和掌握结构化数据、半结构化数据、非结构化数据的概念。
- 6、了解联机分析处理技术（OLAP）的概念、特征。
- 7、熟悉 OLAP 中的多维分析操作，包括钻取、切片和切块、旋转。
- 8、熟悉数据预处理的原因、数据预处理的方法。
- 9、了解数据清洗、数据集成和变换、数据归约的概念与方法，具有应用上述方法进行数据清洗的能力。

第二部分 专业知识

一、大数据技术与应用

（一）数据可视化

- 1、了解 BI（Business Intelligence）的概念及应用。
- 2、了解常见可视化图形（散点图、折线图、饼图、环图、柱状图）的概念，具有初步的可视化图形展示数据的能力。

（二）大数据处理技术

- 1、掌握云计算的概念、云计算的特征、云计算的三层 SPI 模型（IaaS、PaaS、SaaS）。
- 2、了解虚拟化技术的概念、常用虚拟化方法，了解常用虚拟化软件。
- 3、具有应用虚拟化技术搭建虚拟化平台的能力。
- 4、了解分布式计算、高性能计算、边缘计算、工业互联网、5G 的概念。
- 5、熟悉数据集市、数据仓库、数据中台、数据平台的概念。
- 6、熟悉数据采集与预处理过程，掌握数据采集、数据清洗、数据转换、数据脱敏的方法。

7、了解大数据处理环节的主要思想：大数据采集、大数据预处理、大数据存储及管理、大数据分析及挖掘、大数据展现和应用（包括大数据检索、大数据可视化、大数据应用、大数据安全等）。

8、了解大数据处理与分析的代表性流行产品。

二、大数据分析模型

1、了解常用的大数据分析模型，例如行为事件分析模型、点击分析模型、用户行为路径分析模型、用户分群分析模型等，具有应用大数据技术参与大数据分析的能力。

三、数据科学

（一）概率统计与统计推断

- 1、掌握概率、条件概率的概念，并会计算简单的概率、条件概率。
- 2、理解随机事件的独立性和随机变量的独立性。
- 3、掌握乘法公式、全概率公式，并会简单应用。
- 4、掌握随机变量分布函数的概念，掌握连续型随机变量的密度函数和离散型随机变量的分布列。
- 5、掌握常见分布（两点分布、二项分布、泊松分布、均匀分布、正态分布、指数分布）及其简单性质。
- 6、理解数学期望、方差（标准差）、相关系数的概念，并会简单计算。
- 7、理解统计量的概念，理解样本均值、样本方差（标准差）的概念，并会简单计算。
- 8、了解点估计的概念。
- 9、了解线性回归的基本概念。

（二）统计学习与数据挖掘

- 1、掌握特征理解、特征增强、特征构建、特征选择和特征转换等特征工程知识。
- 2、掌握感知机的定义和基本原理。
- 3、掌握 Logistic 回归算法原理与特点，能够使用 Logistic 回归进行数据的分类建模与参数解释。
- 4、掌握朴素贝叶斯算法的基本原理。
- 5、掌握 k 近邻算法的基本原理。
- 6、掌握支持向量机（SVM）算法的基本原理。

- 7、掌握决策树算法的基本原理，了解决策树的剪枝理论。
- 8、了解常见集成方法如 **boosting**、**bagging** 等。
- 9、掌握聚类分析的相关概念。
- 10、掌握主成分分析的基本原理。
- 11、了解生成式半监督学习方法、半监督 **SVM**、图半监督学习方法的基本原理。

（三）深度学习与强化学习

- 1、掌握全连接网络（**MLP**）的定义与算法原理，并能够应用。
- 2、熟悉卷积神经网络（**CNN**）的基本原理，了解几种经典的卷积神经网络，如 **AlexNet**、**VGG**、**GoogLeNet**、**ResNet**。
- 3、熟悉循环神经网络（**RNN**）的基本原理，了解几种常见的模型，如长短期记忆网络 **LSTM**、**GRU**。
- 4、了解马尔可夫决策过程的定义及应用范围。
- 5、了解经典强化学习算法（**Q-learning**、**SARSA**）的基本原理。
- 6、了解经典深度强化学习算法（**DQN**、**DDPG**、**A3C**、**TRPO**、**PPO**）的基本原理，了解其算法特性。